

# Uniqueness of Maronna's $M$ -estimators of scatter

Yacine Chitour, Romain Couillet, and Frédéric Pascal

**Abstract**—In this paper, we prove the uniqueness of Maronna's  $M$ -estimator of scatter [1] for  $N$  vector observations  $c_1, \dots, c_N \in \mathbf{R}^m$  under a mild constraint of linear independence of any subset of  $m$  of these vectors. This entails in particular almost sure uniqueness for random vectors  $c_i$  with a density as long as  $N > m$ . Further relations are established which demonstrate that a properly normalized Tyler's  $M$ -estimator of scatter [2] can be considered as a limit of Maronna's  $M$ -estimator. These results find important implications in recent works on the large dimensional (random matrix) regime of robust  $M$ -estimation.

## I. INTRODUCTION

Subsequent to Huber's introduction of robust statistics in [3], Maronna proposed in [1] a class of robust estimates for scatter matrices defined as the solution of an implicit equation. In [1], the existence and uniqueness of such a solution are proved, under conditions involving both the ratio  $m/N$  of the population size  $m$  and the sample size  $N$ , and the parametrization of the estimate. A sufficient practical condition for uniqueness is that  $m/N$  be small enough. This constraint therefore had little impact on the asymptotic study of these estimators in the regime  $N \rightarrow \infty$  and  $m$  fixed. However, although this constraint was slightly relaxed in [4], [5], with the recent renewed interest in robust  $M$ -estimation under the random matrix regime  $N, m \rightarrow \infty$  with  $m/N \rightarrow c_\infty \in (0, 1)$  [6], [7], [8], [9], Maronna as well as Kent and Tyler conditions have become too stringent and have led these recent works to produce alternative proofs of existence and uniqueness. While Maronna's original results are valid for any (well-behaved) set of samples satisfying the condition on  $m/N$ , the results in e.g. [6] are expressed in probabilistic terms and are only valid for all large  $m, N$ .

Based on the ideas from [10], [11], [12], the present article generalizes both results by showing that existence and uniqueness can be extended to all well-behaved set of samples and for any ratio  $m/N \in (0, 1)$ . In addition,

Chitour is with Laboratoire des Signaux et Systèmes at Supélec, 91192 Gif s/Yvette, France and Université Paris Sud, Orsay, France [yacine.chitour@lss.supelec.fr](mailto:yacine.chitour@lss.supelec.fr). Couillet is with the Telecommunications Department at Supélec [romain.couillet@supélec.fr](mailto:romain.couillet@supélec.fr). Pascal is with the SONDRALaboratory at Supélec [frederic.pascal@supélec.fr](mailto:frederic.pascal@supélec.fr). Couillet's work is funded by ERC-MORE EC-120133.

by a proper parametrization of the weight function, we prove that some sequences of Maronna's  $M$ -estimators converge to Tyler's distribution-free  $M$ -estimator of scatter [2].

The paper is organized as follows: Section II presents our main results, the proofs of which are provided in Section III.

## II. NOTATIONS AND STATEMENT OF THE RESULTS

Let  $m, N$  be positive integers with  $c := \frac{m}{N} \in (0, 1)$ . We use  $M_m(\mathbf{R})$  and  $\text{Sym}_m$  to denote the vector space of  $m \times m$  matrices with real entries and the linear subspace of  $M_m(\mathbf{R})$  made of the symmetric matrices, respectively. We also use  $\text{Sym}_m^+$  and  $\text{PSD}_m$  to denote the non trivial cones in  $M_m(\mathbf{R})$  of the non negative symmetric matrices and of the symmetric positive definite matrices, respectively. Also,  $(\cdot)^T$  stands for the transpose,  $\text{Tr}(\cdot)$  and  $\det(\cdot)$  for the trace and the determinant. On  $M_m(\mathbf{R})$ , we use the inner product defined by the Frobenius norm  $\|A\| = \sqrt{\text{Tr}(AA^T)}$ . We also use  $\leq$  to denote the partial order on  $\text{Sym}_m$  and  $I_m$  the  $m \times m$  identity matrix. Functions of two non negative real variables  $(t, x)$  will be considered. If  $f$  is such a function, we use  $f_t, f_x, f_{tx}, \dots$  to denote (when defined) the partial derivatives of  $f$  with respect to  $t$  and/or  $x$ .

**Definition II.1** A family  $(c_i)_{1 \leq i \leq N}$  of vectors in  $\mathbf{R}^m$  is admissible if

- (C1) for  $1 \leq i \leq N$ ,  $\|c_i\| = 1$ ;
- (C2) for every injective map  $L : \{1, \dots, m\} \rightarrow \{1, \dots, N\}$ , the  $m$  vectors  $c_{L(1)}, \dots, c_{L(m)}$  are linearly independent.

To prove the main result of this paper, the following lemma is required.

**Lemma II.2** Let  $(c_i)_{1 \leq i \leq N}$  be an admissible family of vectors in  $\mathbf{R}^m$ . Fix  $m$  vectors (say)  $c_1, \dots, c_m$  which are then linearly independent by (C2). For  $m+1 \leq l \leq N$ , we can write  $c_l = \sum_{j=1}^m \gamma_{lj} c_j$ . Then,  $\gamma_{lj} \neq 0$  for every  $1 \leq j \leq m$  and  $m+1 \leq l \leq N$ .

**Lemma II.2:** If the conclusion is not true, then there exists an index  $l \geq m+1$  and (at most)  $m-1$  indices in  $\{1, \dots, m\}$  such that the corresponding  $c_i$ 's are linearly dependent, contradicting (C2). ■

We consider maps  $u : (\mathbf{R}_+)^2 \setminus \{(0, 0)\} \rightarrow \mathbf{R}_+$  of class  $C^1$  satisfying:

- (U1)  $u(t, \cdot)$  is strictly decreasing;
- (U2) for every  $t > 0$ ,  $v(t, x) := x \mapsto xu(t, x)$  is increasing on  $\mathbf{R}_+$  and  $l_t := \sup_{x \geq 0} v(t, x) > m$ ;

We furthermore define, for every  $x > 0$ ,  $u(0, x) = \frac{m}{x}$ . Note that, by continuity of  $u$ ,  $\forall x > 0$ ,  $\lim_{t \rightarrow 0^+} v(t, x) = m$ . Also, according to (U2) and (U3), for each  $t, x > 0$ ,

$$v(t, x) = m + tv_1(x) + tw(t, x), \quad (1)$$

with  $v_1(\cdot) := v_t(0, \cdot)$  and  $\forall x > 0$ ,  $\lim_{t \rightarrow 0} w(t, x) = 0$ . By a simple computation, one has that  $v_1$  is a nondecreasing function on  $\mathbf{R}_+^*$ .

For further use, we introduce the following additional notation. Let  $x_t > 0$  be the unique positive number such that,

$$\forall t > 0, v(t, x_t) = x_t u(t, x_t) = m. \quad (2)$$

We further consider the following assumption

$$(U3) \quad \begin{cases} v_x := \partial_x v > 0 \\ v_1 \text{ is increasing} \\ 0 < \liminf_{t \rightarrow 0} x_t \leq \limsup_{t \rightarrow 0} x_t < \infty. \end{cases}$$

If the latter occurs and  $u$  is of class  $C^2$ , then  $w(t, x) = tw_1(x) + o(t)$ , with  $w_1$  continuous on  $(\mathbf{R}_+^*)^2$ , the convergence in (U2) is uniform in  $x$  on any compact of  $\mathbf{R}_+^*$  and  $x_t$  converges to the unique solution  $x_0$  of  $v_1(x) = 0$ .

We use  $\bar{u}(t, x)$  to denote the particular function

$$\bar{u}(t, x) = \frac{m + t}{x + t} \quad (3)$$

which is analytic on every compact of  $(\mathbf{R}_+)^2 \setminus \{(0, 0)\}$ . Moreover,  $l_t = m + t$  and  $v_1(x) = 1 - \frac{m}{x}$ .

The objective of the work is to study the solutions of the equation given, for all  $t > 0$ , by

$$(Eq)_t \quad M = \frac{1}{N} \sum_{i=1}^N u(t, c_i^T M^{-1} c_i) c_i c_i^T.$$

and to characterize them in the limit where  $t \rightarrow 0$ . Taking into account our definitions, if a solution to  $(Eq)_t$  exists, it must belong to  $\text{PSD}_m$ .

Remark that in [1],  $v$  is assumed only to be nondecreasing. Our settings is therefore more constraining in this respect.

To state our results, we need to consider the set of solutions of the equation  $(Eq)_0$  given by

$$(Eq)_0 \quad M = \frac{m}{N} \sum_{i=1}^N \frac{1}{c_i^T M^{-1} c_i} c_i c_i^T.$$

Recall from [10] that the set of solutions of  $(Eq)_0$  is the half-line  $\mathbf{R}_+^* P$  in  $\text{PSD}_m$ , where  $P$  is the unique solution of  $(Eq)_0$  with  $\text{Tr}(P) = m$ .

Our main result is the following theorem.

**Theorem II.3** *Let  $(c_i)_{1 \leq i \leq N}$  be an admissible family of vectors in  $\mathbf{R}^m$  and  $u : (\mathbf{R}_+)^2 \setminus \{(0, 0)\} \rightarrow \mathbf{R}_+$  be a  $C^1$  function verifying (U1)–(U2). Then,*

- (A)  $\forall t > 0$ ,  $(Eq)_t$  admits a unique solution,  $M(t)$ .
- (B) *If, furthermore,  $u$  is  $C^2$  and satisfies (U3), then the mapping  $t \mapsto M(t)$  is continuous and  $\lim_{t \rightarrow 0} M(t) = M_0$  the solution of  $(Eq)_0$  given by  $M_0 = \xi_u P$  with  $\xi_u > 0$  unique solution to*

$$\sum_{i=1}^N v_1 \left( \frac{c_i^T P^{-1} c_i}{\xi} \right) = 0. \quad (4)$$

*In particular, for  $u = \bar{u}$ ,  $M_0 = P$ , i.e.,  $\xi_{\bar{u}} = 1$ .*

To illustrate Theorem II.3, Figure 1 presents the mean square error  $C(t) \triangleq E[\|M(t) - M_0\|_F^2]$  between Tyler's  $M$ -estimator and the Student-t maximum-likelihood estimator (MLE) versus the parameter  $t$  defined through the weight function  $\bar{u}(t, x)$ . We take here  $N = m + 1 = 51$  which ensures, along with the definition of  $\bar{u}(t, x)$  that Maronna's and Kent and Tyler's original conditions are not satisfied. The data are zero-mean Gaussian with Toeplitz covariance matrix, the  $(i, j)$  entry of which is equal to  $\rho^{|i-j|}$ , for some  $\rho \in (0, 1)$ .

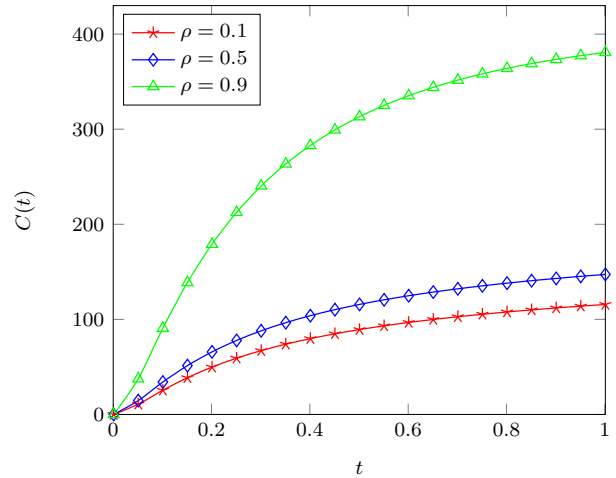


Fig. 1. Convergence of  $M(t)$  towards  $M_0$  when  $t \rightarrow 0$  for  $N = m + 1 = 51$ . The criterion used is the MSE:  $C(t) = E[\|M(t) - M_0\|_F^2]$ .

### III. PROOF OF THEOREM II.3

The strategy of the proof is as follows: for every  $t > 0$ , we first build a positive functional  $H(t, \cdot)$  over  $\text{PSD}_m$  whose critical points (if any) are exactly the solutions of  $(Eq)_t$ . To establish the existence of such critical points, we show that  $H(t, \cdot)$  is uniformly bounded and tends to zero at the boundary of  $\text{PSD}_m$ . To obtain uniqueness, we show that solutions of  $(Eq)_t$  are all local strict maxima

of  $H(t, \cdot)$  and conclude by applying the mountain pass theorem (cf. [13]). This gives Item (A). Item (B) is then obtained using the implicit function theorem and some limiting arguments.

For  $t > 0$ , we define the function

$$\begin{aligned} h : \mathbf{R}_+^* \times \mathbf{R}_+ &\rightarrow \mathbf{R}_+^* \\ (t, x) &\mapsto e^{-\int_{x_t}^x u(t, y) dy}. \end{aligned} \quad (5)$$

Then  $-\frac{h_x}{h} = u$  and  $h(t, x_t) = 1$ . Set  $h(0, x) = \frac{1}{x^m}$  for  $x > 0$  and  $g : \mathbf{R}_+^* \times \mathbf{R}_+ \rightarrow \mathbf{R}_+^*$  with  $g(t, x) = xh(t, x)$ .

In the case where  $u = \bar{u}$ ,  $\forall (t, x) \in (\mathbf{R}_+)^2 \setminus \{(0, 0)\}$ ,

$$x_t \equiv 1, \quad \bar{h}(t, x) = \left(\frac{m+t}{x+t}\right)^{m+t}, \quad \bar{g}(t, x) = x \left(\frac{m+t}{x+t}\right)^{m+t}.$$

Then, define the functional  $H(t, \cdot)$  as

$$\begin{aligned} H : \mathbf{R}_+^* \times \text{PSD}_n &\rightarrow \mathbf{R}_+^* \\ (t, M) &\mapsto \frac{\prod_{i=1}^N h(t, c_i^T M^{-1} c_i)^m}{(\det M)^N} \end{aligned} \quad (6)$$

as well as the functional considered in [10]

$$\begin{aligned} B : \text{PSD}_n &\rightarrow \mathbf{R}_+^* \\ M &\mapsto \frac{\prod_{i=1}^N h(0, c_i^T M^{-1} c_i)^m}{(\det M)^N}. \end{aligned} \quad (7)$$

**Lemma III.1** For  $t > 0$  and  $M \in \text{PSD}_m$ , one has  $-MH_x(t, M)M/NH(t, M) = M - \frac{1}{N} \sum_{i=1}^N u(t, c_i^T M^{-1} c_i)$ , with  $H_x(t, M)$  the gradient of  $H(t, \cdot)$ . In particular,  $M$  is a solution of  $(\text{Eq})_t$  if and only if  $M$  is a critical point of  $H(t, \cdot)$ .

**Lemma III.2**  $\forall t > 0, M \in \text{PSD}_m$ ,  $H(t, M) \leq B(M)$ . As a consequence,  $\lim_{M \rightarrow \partial \text{PSD}_m} H(t, M) = 0$ , so that  $H(t, \cdot)$  admits critical points.

*Lemma III.2:* An immediate calculus yields that  $x \mapsto g(t, x)$  reaches its maximum 1 at  $x = x_t$ . As a consequence, for  $t > 0, M \in \text{PSD}_m$ ,  $H(t, M) \leq B(M)$ . Moreover,  $\lim_{x \rightarrow 0, \infty} g(t, x) = \lim_{x \rightarrow 0, \infty} xh(t, x) = 0$ . For the limit at  $x = 0$ , this is obvious. For  $x \rightarrow \infty$ , note that  $\ln(g(t, x)) = \int_{x_t}^x \frac{m-yu(t, y)}{y} dy$  and, since  $m-l_t < 0$ , it is equivalent to  $(m-l_t) \ln(x)$  as  $x \rightarrow \infty$ . Consider now a sequence  $(M_k)_{k \geq 0}$  in  $\text{PSD}_m$  converging to  $\partial \text{PSD}_m$ . For  $k \geq 0$ , set  $M_k = \rho_k N_k$  with  $\rho_k = \|M_k\|$  and  $N_k = \frac{M_k}{\rho_k}$ . Note that  $\partial \text{PSD}_m$  is made of matrices either non invertible or with norm going to infinity. Therefore, up to subsequences, either (i)  $(N_k)_{k \geq 0}$  converges itself to  $\partial \text{PSD}_m$  or (b) the sequence  $(\rho_k)_{k \geq 0}$  converges to zero or infinity and there exists  $\exists \alpha > 0, \forall k \geq 0, N_k \geq \alpha I_m$ . If Case (i) occurs, then  $\forall k \geq 0, H(t, M_k) \leq B(N_k)$ , which tends to zero as  $k \rightarrow \infty$  (cf. [10]). In Case (ii),

$$H(t, M_k) = \frac{\prod_{i=1}^N h(t, x_{i,k})^m}{\rho_k^N \det(N_k)^N} = B(N_k) \prod_{i=1}^N g(t, x_{i,k})^m$$

where  $x_{i,k} = c_i^T N_k^{-1} c_i / \rho_k$ . As  $k \rightarrow \infty$ ,  $x_{i,k}$  tends either to zero or infinity and we conclude. For  $t > 0$ ,  $H(t, \cdot)$  is uniformly bounded over  $\text{PSD}_m$  since  $B(\cdot)$  is. So  $H(t, \cdot)$  has a global maximum which must belong to  $\text{PSD}_m$  since  $H(t, M) \rightarrow 0$  as  $M$  tends to the boundary of  $\text{PSD}_m$ . So  $H(t, \cdot)$  admits critical points. ■

**Lemma III.3** Let  $t > 0$ . Then all critical points of  $H(t, \cdot)$  are local strict maxima.

*Lemma III.3:* We show that, if  $M$  is a critical point then the Hessian of  $H(t, \cdot)$  at  $M$  is a negative definite quadratic form implying that  $M$  is a local strict maximum of  $H(t, \cdot)$ . Let  $M \in \text{PSD}_m$  be a critical point of  $H(t, \cdot)$ . Then, one gets that for every  $Q \in \text{Sym}_m$ ,

$$\begin{aligned} \langle Q, \text{Hess}_M(Q) \rangle &= -NH(t, M) \left[ \langle Q, M^{-1} Q M^{-1} \rangle \right. \\ &\quad \left. + \frac{1}{N} \sum_{i=1}^N u_x(t, c_i^T M^{-1} c_i) (c_i^T M^{-1} Q M^{-1} c_i)^2 \right]. \end{aligned}$$

With  $R := M^{-1/2} Q M^{-1/2}$  and  $d_i := M^{-1/2} c_i$ , we have

$$\begin{aligned} -\frac{\langle Q, \text{Hess}_M(Q) \rangle}{NH(t, M)} &= \\ \|R\|^2 + \frac{1}{N} \sum_{i=1}^N u_x(t, \|d_i\|^2) (d_i^T R d_i)^2. \end{aligned} \quad (8)$$

Recall that  $M$  is a critical point of  $H(t, \cdot)$  and thus a solution of  $(\text{Eq})_t$ , i.e.,

$$I_m = \frac{1}{N} \sum_{i=1}^N u(t, \|d_i\|^2) d_i d_i^T. \quad (9)$$

Multiplying (9) by  $R$  on both left and right, taking the trace and plugging the result into (8) gives

$$(8) = \frac{1}{N} \sum_{i=1}^N u(t, \|d_i\|^2) \|R d_i\|^2 + u_x(t, \|d_i\|^2) (d_i^T R d_i)^2.$$

Let  $I_Q = \{i \in \{1, \dots, N\}, R d_i \neq 0\}$ . Then

$$(8) = \frac{1}{N} \sum_{i \in I_Q} \|R d_i\|^2 [u(t, \|d_i\|^2) + \|d_i\|^2 u_x(t, \|d_i\|^2) r_i]$$

where  $r_i := (d_i^T R d_i / [\|d_i\| \|R d_i\|])^2$ . Using  $0 \leq r_i \leq 1$  (by Cauchy-Schwarz's inequality) and  $u_x \leq 0$  (since  $u$  is of class  $C^1$  and verifies (U1)), we have  $r_i u_x(\cdot, \cdot) \geq u_x(\cdot, \cdot)$ . Then, recalling that  $v(t, x) = x u(t, x)$ ,

$$(8) \geq \frac{1}{N} \sum_{i \in I_Q} \|R d_i\|^2 v(t, \|d_i\|^2) \geq 0.$$

Moreover, if  $Q \neq 0$ ,  $I_Q \neq \emptyset$  and there exists  $\bar{i}$  such that  $v_x(t, \|d_{\bar{i}}\|^2) > 0$ . Therefore  $\langle Q, \text{Hess}_M(Q) \rangle < 0$ , i.e.,  $\text{Hess}_M$  is negative definite, concluding the proof. ■

**Lemma III.4** Let  $t > 0$ . Then  $(\text{Eq})_t$  admits a unique solution,  $M(t)$ , the unique strict maximum of  $H(t, \cdot)$ .

**Lemma III.4:** We reason by contradiction assuming  $H(t, \cdot)$  admits at least two local strict maxima. Applying the mountain-pass theorem [13] to the functional  $1/H(t, \cdot)$  which tends to infinity in the vicinity of  $\partial \text{PSD}_m$ , we obtain the existence of a saddle point of  $F$  in  $\text{PSD}_m$  which is contradictory to Lemma III.3. ■

We next prove that  $M(t)$  is uniformly bounded in  $\text{PSD}_m$  as  $t \rightarrow 0$ , i.e.,

**Lemma III.5** *There exists  $0 < a \leq b$  and  $t_0 > 0$  such that, for every  $t \in (0, t_0)$ ,  $aI_m \leq M(t) \leq bI_m$ .*

**Lemma III.5:** Let  $P$  be the unique matrix of  $\text{PSD}_m$  satisfying  $B(P) = \max_{M \in \text{PSD}_m} B(M)$  and  $\text{Tr}(M) = m$ . Then, for every  $t > 0$ ,  $H(t, P) \leq H(t, M(t))$  and  $B(M(t)) \leq B(P)$ . Multiplying both inequalities, after simplifications, we get  $\prod_{i=1}^N g(t, c_i^T P^{-1} c_i) \leq \prod_{i=1}^N g(t, c_i^T M(t)^{-1} c_i) \leq 1$ , with  $\prod_{i=1}^N g(t, c_i^T P^{-1} c_i) \rightarrow 1$  as  $t \rightarrow 0$ . So there exists  $t_0 > 0$  such that, for every  $t \in (0, t_0)$  and  $1 \leq i \leq N$ ,  $1/2 \leq g(t, c_i^T M(t)^{-1} c_i)$ , and, since (U3) holds true, there exists  $0 < a \leq b$  such that, for every  $t \in (0, t_0)$  and  $1 \leq i \leq N$ ,  $a \leq c_i^T M(t)^{-1} c_i \leq b$ . This implies that, for every  $t \in (0, t_0)$  and  $1 \leq i \leq N$ ,  $u(t, b) \leq u(t, c_i^T M(t)^{-1} c_i) \leq u(t, a)$ , hence  $u(t, b)C \leq M(t) \leq u(t, a)C$  where  $C := \frac{m}{N} \sum_{i=1}^N c_i c_i^T$ . One concludes easily. ■

**Lemma III.6** *Under the conditions of Theorem II.3,  $\lim_{t \rightarrow 0} M(t) = M_0$  solution of  $(\text{Eq})_0$  given by  $M_0 = \xi_u P$ , where  $\xi_u > 0$  is the unique solution of (4).*

**Lemma III.6:** Since  $M(\cdot)$  is uniformly bounded in  $\text{PSD}_m$  as  $t \rightarrow 0$ , its accumulation points still belong to  $\text{PSD}_m$  and are necessarily of the form  $\mu P$  where  $\mu > 0$  and  $P$  is the solution of  $(\text{Eq})_0$  with trace  $m$ . Taking the trace in (9), one gets  $m = \frac{1}{N} \sum_{i=1}^N v(t, \|d_i(t)\|^2)$ , where  $d_i(t) = M(t)^{-1/2} c_i$  for  $1 \leq i \leq N$ . Using (1) and (U3), one deduces that, for every  $t > 0$ ,  $\sum_{i=1}^N v_1(\|d_i(t)\|^2) + t \sum_{i=1}^N w_1(\|d_i(t)\|^2) + o(t) = 0$ . Consider an accumulation point  $\mu P$  of  $M(\cdot)$  as  $t \rightarrow 0$ . Then, up to a subsequence,  $\lim_{t \rightarrow 0} M(t) = \mu P$  and, for  $1 \leq i \leq N$ ,  $\lim_{t \rightarrow 0} d_i(t) = P^{-1/2} c_i / \sqrt{\mu}$ . According to (U3), the second sum in the previous equation tends to zero as  $t \rightarrow 0$  and we are left with  $\sum_{i=1}^N v_1(c_i^T P^{-1} c_i / \mu) = 0$ . Since the left-hand side of this equation defines a decreasing function of  $\mu$ , it has a unique solution denoted  $\xi_u > 0$ , which concludes the proof since  $M(\cdot)$  admits a unique accumulation point as  $t \rightarrow 0$ . ■

#### IV. CONCLUSIONS

In this paper, we relaxed the conditions given in [1] to prove the existence and uniqueness of scatter matrix

$M$ -estimators of the Maronna type. Uniqueness is in particular ensured for all population size  $m$  and sample size  $N$  as long as  $m < N$  and sample vectors satisfy a linear independence condition. This result is of important interest in application contexts where  $m$  and  $N$  are of the same order of magnitude, as opposed to classical robust statistics where it is traditionally assumed that  $N \gg m$ . This is of particular importance for recent studies on robust statistics based on random matrix theory for which  $m$  and  $N$  are both large but with non trivial ratio. Furthermore, we drew some connections between Maronna's and Tyler's estimators by expressing (properly scaled) Tyler's estimator in terms of a limit of a generic class of Maronna's estimators. This result may also find interest in studies of Tyler's  $M$ -estimator in the large random matrix regime.

#### REFERENCES

- [1] R. A. Maronna, "Robust  $M$ -estimators of multivariate location and scatter," *Annals of Statistics*, vol. 4, no. 1, pp. 51–67, January 1976.
- [2] D. Tyler, "A distribution-free  $m$ -estimator of multivariate scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.
- [3] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.
- [4] D. Tyler, "Some results on the existence, uniqueness, and computation of the  $M$ -estimates of multivariate location and scatter," *SIAM Journal on Scientific and Statistical Computing*, vol. 9, p. 354, 1988.
- [5] J. T. Kent and D. E. Tyler, "Redescending  $M$ -estimates of multivariate location and scatter," *Annals of Statistics*, vol. 19, no. 4, pp. 2102–2119, December 1991.
- [6] R. Couillet, F. Pascal, and J. Silverstein, "Robust  $M$ -Estimation for Array Processing: A Random Matrix Approach," *Information Theory, IEEE Transactions on* (submitted), 2013. [Online]. Available: arXiv:1204.5320. [Online]. Available: http://arxiv.org/abs/1204.5320
- [7] R. Couillet, F. Pascal, and J. W. Silverstein, "The Random Matrix Regime of Maronna's  $M$ -estimator with elliptically distributed samples," *Journal of Multivariate Analysis* (submitted), 2013. [Online]. Available: arXiv:1311.7034.
- [8] T. Zhang, X. Cheng, and A. Singer, "Marchenko-Pastur Law for Tyler's and Maronna's  $M$ -estimators," *arXiv preprint arXiv:1401.3424*, 2014.
- [9] I. Solovychik and A. Wiesel, "Non-asymptotic Error Analysis of Tyler's Scatter Estimator," *arXiv preprint arXiv:1401.6926*, 2014.
- [10] F. Pascal, Y. Chitour, J. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: existence and algorithm analysis," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 34–48, Jan. 2008.
- [11] Y. Chitour and F. Pascal, "Exact maximum likelihood estimates for SIRV covariance matrix: existence and algorithm analysis," *Signal Processing, IEEE Transactions on*, vol. 56, no. 10, pp. 4563–4573, Oct. 2008.
- [12] F. Pascal, Y. Chitour, and Y. Quek, "Generalized robust shrinkage estimator and its application to STAP detection problem," *Signal Processing, IEEE Transactions on* (submitted), 2013. [Online]. Available: arXiv:1311.6567.
- [13] M. Struwe, *Variational methods: applications to nonlinear partial differential equations and Hamiltonian systems*, 4th ed., Springer, Ed. Springer, 2008, vol. 34.